
DÉVELOPPEMENT D'UNE ÉVALUATION GÉNOMIQUE POUR L'ANALYSE DE DONNÉES LONGITUDINALES : APPLICATION AUX CONTRÔLES ÉLÉMENTAIRES CHEZ LES CAPRINS

Thèse de Mathieu **ARNAL**¹

Analysée par Jean-Claude **MOCQUOT**²

Directrice de thèse : Mme Christèle Robert-Granié, Directeur de recherche INRAE

Co-directrice de thèse : Mme Hélène Larroque, Ingénieur de recherche INRAE

L'objet de cette thèse est de conduire les recherches nécessaires à la mise en pratique, dans l'espèce caprine, d'une évaluation génétique et génomique des reproducteurs, à partir de chacun des contrôles laitiers réalisés au cours de la lactation, et non plus à partir de la production totale de chaque lactation estimée, à partir de ces contrôles par intrapolation et cumul de ces données élémentaires.

Cette méthodologie d'analyse de données longitudinales, déjà mise en œuvre pour les bovins laitiers, présente plusieurs intérêts : outre une meilleure prise en compte des effets d'environnement affectant la production le jour de chaque contrôle, elle permet la modélisation de la plupart des effets génétiques et d'environnement sous forme de courbes. Elle offre, en particulier, la possibilité d'une évaluation de la forme de la courbe de lactation, donc une action de sélection potentielle sur le caractère de persistance des lactations.

Une telle action sur la forme des courbes de lactation est notoirement importante pour les caprins, d'abord, parce que, comme pour les bovins, l'augmentation forte des productions s'est accompagnée d'un pic de plus en plus élevé en début de lactation, avec les difficultés et coûts d'alimentation et problèmes de troubles métaboliques associés ; ensuite l'étalement de la production par le biais d'une persistance accrue est un atout global pour cette espèce dont les mises-bas sont très saisonnées, alors que les attentes du marché et de la transformation fromagère ne le sont pas.

Ce travail s'est déroulé en cinq parties successives résumées ci-après.

I - INTRODUCTION GENERALE

En une trentaine de pages, l'auteur décrit de façon très exhaustive et claire la filière caprine française, l'organisation de la collecte des données, l'organisation de la sélection, la méthodologie d'évaluation génétique pour les différents caractères laitiers et les décomptes de cellules somatiques, ainsi que les paramètres génétiques calculés pour les races Saanen et Alpine.

Les dix dernières pages présentent les limites de ces évaluations génétiques à partir de lactations totales, et présente les différentes méthodologies d'approche des évaluations réalisées et mises en œuvre en bovins laitiers à partir des contrôles élémentaires, dans de nombreux pays.

Cette partie bibliographique est mise à profit pour orienter et justifier les choix méthodologiques qui seront testés et/ou retenus dans la suite de la thèse pour la situation particulière des caprins.

II - ÉTUDE DE LA DIVERSITÉ DES COURBES DE LACTATION DES CHÈVRES ALPINE ET SAANEN

Cette étude phénotypique, préalable à l'étude des facteurs génétiques et environnementaux de variation de la forme des courbes de lactation, a fait l'objet d'une publication en 2018 dans le « *Journal of Dairy Science* », intégrée dans la thèse.

Le travail a consisté en une analyse en composantes principales (ACP) des données de plus de deux millions de contrôles élémentaires, représentant 320 000 lactations de 213 534 chèvres.

Les trois premières composantes représentaient 99,5 % de la variance. Le premier axe (81%) rendait compte du niveau de production sur l'ensemble de la lactation, le second (15%) traduisait la persistance et le troisième (3,5%) rendait compte du niveau de production en milieu de lactation.

Une méthodologie appropriée de classification à partir des valeurs des trois composantes principales a ensuite permis de valider des clusters de formes de courbes de lactations. Le choix de cinq clusters a été retenu comme le plus pertinent après des approches complémentaires, également présentées en sus de la publication, qui faisaient varier le nombre de clusters entre trois et huit.

Les deux premiers clusters, représentant 70 % des lactations, se différençaient par le niveau global de production, avec des formes de courbe proches de la courbe moyenne. Les trois autres clusters, représentant 30% des lactations, présentaient des courbes de forme très différentes, l'un avec une très faible persistance, les deux autres avec de très bonnes persistances, soit pendant toute la lactation, soit pendant la moitié de la lactation. Une analyse des facteurs de variation de cette diversité de courbes a notamment révélé l'influence prépondérante du mois de mise-bas sur chacune des trois composantes, mais aussi des différences entre régions, races, stades de gestation et durées du tarissement précédent.

Pour terminer ce chapitre l'auteur a présenté graphiquement l'évolution des courbes de lactation moyennes au cours des neuf dernières années, pour les caractères contrôlés (lait, TB, TP, TB/TP, MG, MP, cellules somatiques) selon le rang de lactation et la race.

III - ESTIMATION DES PARAMÈTRES GÉNÉTIQUES A L'AIDE D'UN MODÈLE DE RÉGRESSION ALÉATOIRE

Cette partie originale, et plus proprement génétique, est présentée sous forme de deux séries de travaux et d'études complémentaires de tests de validation des choix méthodologiques retenus ou des hypothèses faites. Elle est trop spécialisée en méthodologie statistique pour en résumer le contenu. Elle révèle une excellente maîtrise par l'auteur de ces aspects très pointus. Nous n'évoquerons ici que les données et résultats.

Le premier article publié en 2019 dans « *Genetic Selection Evolution* » estime les paramètres génétiques en première lactation, à partir des contrôles élémentaires ajustés par un modèle de régression aléatoire, pour les races Saanen et Alpine. Cinq modèles d'ajustement de données sont comparés pour rechercher celui qui garde une pertinence suffisante pour des temps de calculs et ressources de mémoire plus faibles. Les résultats sont comparés avec ceux issus du modèle classique sur les données de production totale par lactation. De ce nombre impressionnant de calculs et de résultats, il est apparu qu'un des modèles réduits prenant en compte les deux premières composantes de l'ACP assurait à la fois des calculs praticables, y compris avec les lactations de rang supérieur à un, une bonne pertinence et une interprétation claire des deux caractères de production et de persistance. Les héritabilités sont très voisines avec chacun des modèles d'ajustement, de l'ordre de 0,3 à 0,37 pour les quantités (Lait, MP, MG), 0,68 à 0,75 pour les taux, et 0,18 pour la persistance laitière. Par ailleurs, quantité et persistance sont en corrélation génétique quasi nulle, ce qui est un résultat très intéressant pour une sélection combinée des deux caractères. Le second article soumis pour publication dans « *Journal of Dairy Science* » concerne l'estimation des paramètres génétiques des mêmes sept caractères au cours des trois premières lactations avec la même méthodologie. Quatre modèles ont été comparés pour la qualité de l'ajustement aux données, l'évolution des héritabilités et de la variance génétique au cours et entre rangs de lactation.

Les corrélations génétiques entre stades de lactation des différentes parités, et les liaisons entre les valeurs génétiques estimées (Index), ont aussi été analysées. En complément de l'article, les effets de milieu issus du modèle, les variances génétiques, ainsi qu'une première approche des liaisons entre persistance et les caractères de fertilité à l'IA, de la survie productive ou longévité, ont été analysées.

L'ensemble de ces travaux permet d'envisager concrètement une évaluation génétique de routine à partir de ces données longitudinales.

IV - ÉTUDE DE LA FAISABILITÉ D'UNE ÉVALUATION GÉNOMIQUE BASÉE SUR LES CONTRÔLES ÉLÉMENTAIRES

Une évaluation génomique à partir des résultats de la puce de SNP 50K est en pratique depuis 2018 pour les lactations totales des caprins avec un modèle Single step GBLUP. L'auteur a donc proposé d'appliquer cette méthodologie aux contrôles élémentaires modélisés comme dans le volet précédent par régression aléatoire.

La validation de cette évaluation génomique sur contrôles élémentaires a fait l'objet d'un article soumis au « *Journal of Dairy Science* », la méthodologie mise en œuvre a permis de comparer les résultats obtenus avec ceux basés sur les lactations totales et aussi avec les évaluations classiques des boucs sur descendance.

Les résultats sont globalement très concordants, l'évaluation génomique sur contrôles élémentaires s'avère praticable en routine, réduit certains biais, améliore la corrélation avec les évaluations sur descendance et surtout permet une évaluation du caractère de persistance.

Une analyse complémentaire des effets des SNP a permis de préciser les régions chromosomiques plus associées à chacun des caractères évalués et certaines sur les chromosomes 6 et 19 semblent notamment impliquées dans la variation de la persistance des lactations.

La discussion générale récapitule tous les apports de ces travaux, tant pour la qualité d'évaluation des différents caractères et de leurs liaisons génétiques, que pour l'utilisation possible des estimées des facteurs d'environnement pour l'appui technique aux éleveurs.

La qualité et la quantité de travail accompli lors de cette thèse sont très remarquables et impossibles à résumer brièvement. Les quatre publications dans les revues internationales attestent de cette excellence. Par ailleurs ces travaux ouvrent une voie nouvelle et immédiatement applicable à la sélection des caprins français pour mieux répondre aux attentes de cette filière.

Cet important travail scientifique mérite que cette analyse de thèse figure sur le site de l'Académie à titre de valorisation.